PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 7/01, 15/12, 15/70

A1

(11) International Publication Number: WO 96/21006

(43) International Publication Date: 11 July 1996 (11.07.96)

(21) International Application Number: PCT/US95/15796

(22) International Filing Date: 6 December 1995 (06.12.95)

(30) Priority Data:
08/369,494 6 January 1995 (06.01.95) US

(71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd floor, 300 Lakeside Drive, Los Angeles, CA 94612-3550 (US).

(72) Inventor: PRICE, Paul, A.; 8695 Cliffridge Avenue, La Jolla, CA 92093 (US).

(74) Agent: BERLINER, Robert; Robbins, Berliner & Carson, 5th floor, 201 N. Figueroa Street, Los Angeles, CA 90012-2628 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: NOVEL PHOSPHOPROTEIN SECRETED IN THE EXTRACELLULAR MATRICES OF MAMMALIAN ORGANS AND METHODS FOR USE THEREOF

(57) Abstract

A phosphoprotein component of mammalian connective tissue (particularly bone; "Spp24") has been identified, cloned, sequenced and expressed in recombinant expression vectors. Spp24 protein is highly homologous to cystatin domain, cysteine protease inhibitors which are involved in regulation of connective tissue metabolism in several mammalian species. Antibodies, Spp24 and compositions containing the antibodies or Spp24 are provided. Methods for use of Spp24 and antibodies thereto in therapeutic, diagnostic and research protocols are also described.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

NOVEL PHOSPHOPROTEIN SECRETED IN THE EXTRACELLULAR MATRICES OF MAMMALIAN ORGANS AND METHODS FOR USE THEREOF

BACKGROUND OF THE INVENTION

1. Field of the Invention.

The invention relates to a novel phosphoprotein which plays a role in the resorption and formation of extracellular matrices in mammalian organs, particularly in bone. The invention further relates to methods for the use of the novel phosphoprotein in the investigation and control of such events.

2. History of the Prior Art.

The extracellular matrix of mammalian connective tissues (such as the mineralized collagen fibers of bone, the articular cartilage of joints, and the vascular wall tissue of the vascular and lymphatic system) provides a barrier to the migration of cells from the tissue and strength to it. Of these tissues, bone in particular is unusual in that the extracellular matrix therein is constantly being remodeled through a process of resorption (mediated by osteoclasts) and formation (mediated by osteoblasts) of the matrix. This process is generally depicted in FIGURE 1.

The precise mechanism by which osteoblasts and osteoclasts interact in bone remodeling is unknown. In general, it is known that proteins secreted by such cells serve to mediate the remodeling process, including remineralization and release of calcium from the extracellular matrix to maintain serum calcium homostasis. However, to date, only a few constituent bone matrix proteins have been isolated and identified; of these, relatively little is known about their function in matrix turnover. These proteins principally include: bone Gla protein (Price, *et al., Proc. Nat'l. Acad. Sci. USA, 73:3374-3375*, 1976; Pan and Price, *Proc. Nat'l. Acad. Sci. USA*, 82:6109-6113, 1985);

matrix Gla protein (Price and Williamson, *J. Biol. Chem.*, 260:14971-14972, 1985); osteoponin (Oldberg, *et al., Proc. Natl. Acad. Sci USA*, 83:8319, 1986); bone sialoprotein (Oldberg, *et al., J. Biol. Chem.*, 263:19430, 1988); and, the proteoglycans decorin, biglycan and osteonectin (*see.* e.g., Day, *et al., Biochem. J.*, 248:801, 1987; and, Mason, *et al., EMBO J.*, 5:1831, 1986).

Bone resorption is believed to be enzymatically regulated by two classes of proteases: the metalloproteases (such as collagenase) and the cysteine proteases (such as the cathepsins and beta-cystatin). The activity of these proteases is modulated in different species by naturally occurring inhibitors, such as alpha-2-HS-glycoprotein in humans, fetuin in bovines and rats, PLCPI in pigs, and human cystatin C in humans, as well as being regulated by calcitonin (as tested in humans and mice). One recent model for the activity of all of these proteins proposes that the cysteine proteases are principally involved in bone resorption (on activation through disruption of a zinc-cysteine bond in the enzyme), while metalloproteases are secondarily responsible for the "pit formation" which characterizes the bone resorption process (*see*, e.g., Everts, *et al., J. Cell Physiol.*, 150:221-231, 1992).

Many of the known cysteine protease inhibitors share common cystatin domains, even across species boundaries. To some extent, this homology is not unexpected. Given the observation that mammalian osteoinductive proteins tend to share a relatively high degree of homology (*see*, e.g., Deatherage, *et al., J. Oral Maxillofac. Surg.*, 17:395-399, 1988), it would not be surprising if many of the proteases which degrade such proteins, and inhibitors which retard the activity of such proteases, would each share common primary structures with other proteases and inhibitors.

Usually, the resorption and reformation of extracellular matrices is beneficial to the host, permitting normal metabolism, growth and tissue repair to occur. However, in certain disease processes (e.g., osteoarthritis) the matrix is excessively degraded by hydrolytic proteases (e.g., collagenases). As the matrix degrades, the integrity of the tissue is impaired, which may allow tissue cells, by-products and other residues of the matrix

metabolism (such as inhibitory molecules) to escape into bodily fluids and/or lymphatic or vascular circulation.

Detection of these molecules can, in certain instances, provide information regarding the biochemical characteristics of the extracellular matrix, including how it is synthesized and how it is lost. Also, where a particular molecule that is produced and/or secreted during abnormal matrix metabolism is closely related to a disease process, quantification of that molecule in the patient's body fluids and/or tissues can help clinicians track the progress of, as well as treat, the disease.

Conversely, in certain pathologies, matrix metabolism is subnormal. For example, the osteopetrosis-like bone disorder pycnodysostosis is characterized by insufficient osteoclastic-mediated resorption activity. The pathological effects of this disorder can be mimicked *in vitro* by addition of an excess of cysteine protease inhibitors to cultured bone explants (*see*, e.g., Everts, *et al.*, *Calcif. Tissue Int.*, 43:172-178, 1988).

Even when matrix metabolism proceeds normally, knowledge of the structure and function of molecules which regulate that metabolism can be used to prophylactically enhance matrix metabolism to benefit the host. For example, devices such as the ones described in U.S. Patent No. 5,041,188 to Vacanti, *et al.* are utilized *in vivo* as artificial matrices to encourage formation of bone or cartilage at specific locations to, for example, ensure adequate repair of injured bone or cartilage tissue. It has been shown that the efficacy of such devices can be enhanced by the addition of native matrix constituents to the artificial matrix environment (*see*, e.g., Deatherage, *et al.*, *J. Oral Maxillofac. Surg.*, 17:395-399, 1988).

Thus, identification of other protease inhibitors active in bone and other matrix-intensive organs would be of substantial value in research and development relating to control of extracellular matrix metabolism.

SUMMARY OF THE INVENTION

The invention is based on the discovery of a novel phosphoprotein which possess a degree of sequence homology, structural characteristics and tissue distribution consistent with those of cysteine protease inhibitors found in bone and liver tissue.

In one aspect, the invention comprises the protein and peptides derived therefrom.

In another aspect, the invention comprises assays for homologous proteins and peptides in other species and for detection of the novel protein in body tissues or fluids.

In another aspect, the invention comprises methods for enhancing or retarding the activity of cysteine-proteases through interaction with the novel protein.

In another aspect, the invention comprises devices for enhancing or retarding the activity of cysteine-proteases through interaction with the novel protein.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 depicts the general processes associated with bone formation and resorption in vertebrates.

FIGURE 2. Purification of Spp24 and MGP by gel filtration over a Sephacryl S-100 HR column. The proteins extracted from ground bovine bone by demineralization in 10% formic acid were dried and the neutral pH-soluble proteins were removed by washing repeatedly with 50 mM $NH_4H_2O_3$. The water-insoluble proteins were then dissolved in 6 M guanidine HCl with 0.1 M Tris HCl at pH 9.0 and loaded onto a 2 x 150 cm Sephacryl S-100 HR column equilibrated with the same buffer at room temperature. Fraction volume, 3 ml.

Inset: SDS-polyacrylamide gel electrophoresis of partially purified Spp24. Proteins were electrophoresed on a 4-20% gradient gel and stained with Coomassie Brilliant Blue. Lane 1, molecular weight standards; lane 2, 20 μg of the purified Spp24 in pooled fractions 61-3.

FIGURE 3. Further purification of Spp24 by reverse phase high pressure liquid chromatography. 100 μg of the partially purified Spp24 in pooled fractions 61-3 from the gel filtration shown in Figure was loaded directly onto a 4.6mm x 25cm C_4 column equilibrated with 0.1% trifluoroacetic acid at room temperature. Bound proteins were subsequently eluted with a 2 h linear gradient to 0.1% trifluoroacetic acid in 60% acetonitrile. Fraction volume, 1.3 ml.

Inset: SDS-polyacrylamide gel electrophoresis of purified Spp24. Proteins were electrophoresed on a 4-20% gradient gel and stained with Coomassie Brilliant Blue. Lane 1, molecular weight standards; lane 2, 10 μg of the purified Spp24 in pooled fractions 24-27.

FIGURE 4. Strategy for determining the sequence of Spp24 cDNA. (1) A 380-bp cDNA fragment which contains the N-terminal portion of Spp24 was obtained using RT-PCR with two degenerate primers. (2) A 370-bp cDNA fragment which contains the 3'-untranslated region of Spp24 was generated using 3'-RACE with a specific internal primer. (3) A 312-bp fragment of Spp24 was used for λgt11 cDNA library screening and Northern blot analysis. (4) A 470-bp cDNA fragment was obtained by screening a bovine liver λgt11 cDNA library with the 312-bp cDNA probe and the 3'-region (243-bp) of this fragment was sequenced. The primers used in this regard are shown in SEQ. ID NOS: 1 through 6.

FIGURE 5. Complete nucleotide sequence of Spp24 cDNA and deduced amino acid sequence of the protein (SEQ ID NOS:7 and 8). *Underlined* sequences correspond to those determined by protein sequencing. The cleavage site of the putative signal sequence is indicated by the *arrow* at residue 20 and the N-terminus of mature Spp24 begins at residue 21. The stop codons are marked by *asterisks*. The polyadenylation signal sequence AATAAA is indicated in *bold lettering*.

FIGURE 6. Northern blot analysis of Spp24 message levels in bovine tissues. RNA was extracted from the indicated bovine tissues and 40 μg of total RNA from each tissue was run on a 1.4% formaldehyde-agarose gel, blotted onto a Nytran™ membrane, and hybridized with a ^{32}p-labeled, 312-bp Spp24 cDNA fragment (FIGURE 5). Lane 1, bone periosteum; Lane 2, heart; Lane 3, lung; Lane 4, kidney; Lane 5, spleen; Lane 6, liver. The migration positions of molecular size markers in kilobases are indicated on the *left*.

FIGURE 7. Amino acid sequence homologies between Spp24 and porcine cathelin, bovine bactenecin precursor, cystatin domains 1 and 3 of human kininogen, and chicken cystatin. Residue numbers refer to the sequence position in mature Spp24. The related sequences are residues 1-92 of cathelin; 24-126 of bactenecin precursor; 23-127 of kininogen (cystatin domain 1), 268-371 of kininogen (cystatin domain 3), and 12-116 of chicken cystatin. Identical amino acids are boxed.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

A. Spp24 Encoding Polynucleotides and Methods for *In Vitro* Use of Same.

E.coli expression vectors containing, in tandem, a cDNA encoding bovine Spp24 (SEQ.ID.No.7) were deposited with the American Type Culture Collection (ATCC), on January 5, 1995 as Spp24-mid ATCC Designation 97012 (containing a plasmid containing a 376 bp insert of nucleotides 151 through 527 of Spp24 [*see*, FIGURE 5]); Spp24-3' ATCC Designation 97011 (containing a plasmid containing a 370 bp insert of nucleotides 459 through 829 of Spp24 [*id.*]; and, Spp24-5' ATCC Designation 97013 (containing a plasmid containing a 380 bp insert from the 5' untranslated end of Spp24 through to nucleotide 243 [*id.*]).

In addition, a deposit of λgt-11 phage including a 800bp cDNA encoding human Spp24 was made on January 5, 1995. The human Spp24 clone ("h-Spp24") was obtained as described in Example IV; i.e., the Spp24-mid nucleotide was labelled and used to probe a human liver cDNA library. The polynucleotide contained in the human Spp24 phage vector deposit was identified as being the full-length nucleotide sequence coding for Spp24 in human tissues.

No admission that these deposits were necessary to enablement of the claims herein is made or intended. The deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of viable vectors for 30 years from the date of deposit. The vectors will be made available by the ATCC under the terms of the Budapest Treaty which assures unrestricted availability of the vectors to the public upon issuance of the pertinent U.S. patent or open laying open to the public of and U.S. or foreign patent application, whichever comes first, and assures availability of the vectors to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 U.S.C. § 122 and the Commissioner's rules made pursuant thereto.

The invention encompasses polynucleotides which encode Spp24 (including h-Spp24) and the Spp24 peptides of the invention (*see*, e.g., SEQ. ID. Nos. 7 and 8 and ATCC Designation Nos. 97011, 97012 and 97013. As used herein, "polynucleotide" refers to a polymer of deoxyribonucleotides or ribonucleotides, in the form of a separate fragment or as a component of a larger construct. Further, unless context otherwise requires, as used herein, "Spp-24" will include polynucleotides encoding the bovine and human Spp24 proteins, as well as homologs thereof from other mammalian species. DNA encoding a Spp24 peptide of the invention can be assembled from cDNA fragments or from oligonucleotides which provide a synthetic gene which is capable of being expressed in a recombinant transcriptional unit. Polynucleotide sequences of the invention include DNA, RNA and cDNA sequences. A polynucleotide sequence can be deduced from the genetic code, however, the degeneracy of the code must be taken into account. Polynucleotides of the invention include sequences which are degenerate as a result of the genetic code.

Specific DNA sequences encoding Spp24, fragments thereof, or Spp24-homologous peptides can be obtained by: 1) isolation of double-stranded DNA sequences from the genomic DNA: 2) chemical manufacture of a DNA sequence to provide the necessary codons for the polypeptide of interest; and 3) *in vitro* synthesis of a double-stranded DNA sequence by reverse transcription of mRNA isolated from a eukaryotic donor cell. These procedures are conventional ones which will be familiar to those of ordinary skill in the art.

Polynucleotide sequences encoding the peptides of the invention can be expressed in either eukaryotes or prokaryotes, including the transformed *E. coli* and phage vectors which are described above (ATCC Designation Nos. 97011, 97012 and 97013). Hosts can include microbial, yeast, insect and mammalian organisms. Methods of expressing DNA sequences having eukaryotic or viral sequences in prokaryotes are well known in the art. Biologically functional viral and plasmid DNA vectors capable of expression and replication in a host are known in the art. Such vectors are used to incorporate DNA of the invention.

The polynucleotides of the invention are also useful as probes to identify homologous proteins in the tissues of non-bovine species. As shown in FIGURE 7, Spp24 protein is highly homologous in amino acid sequence to other cystatin domain containing protease inhibitors. In addition, as discussed in the Background section, it may be reasonably predicted that substantial homology will exist among any given set of osteoregulatory proteins, given the degree of homology which has been identified in such groups to date. Thus, accounting for degeneracy in the code, the polynucleotides of the invention (in particular those regions which encode the cystatin domain of Spp24; see, e.g., FIGURES 5 and 7; SEQ. ID. No. 7; and, ATCC Designation Nos. 97011, 97012 and 97013, in particular Spp24-mid and h-Spp24) can be expected to be effective probes for identification of homologous proteins. For example, the respective termini of the polynucleotide of SEQ. ID. No. 7 (obtained by, for example, splitting the coding region of the polynucleotide) or cystatin regions thereof (see, e.g., Spp24-mid; ATCC Designation No. 97012) may be effectively used as hybridization probes for complementary sequences in a cDNA library (such as the human cDNA library screened to obtain h-Spp24) or other suitable source material.

More specifically, the polynucleotide probes and antibodies of the invention (the latter of which are described below) can be utilized in this respect in hybridization protocols which are well known in the art. These include, but are not limited to: 1) hybridization of probes to genomic or cDNA libraries to detect shared nucleotide sequences; 2) antibody screening of expression libraries to detect shared structural features and 3) detection by the polymerase chain reaction (PCR). Such protocols are standard in the art and will therefore be only briefly described below. For review, however, those of skill in the art may wish to refer to Ausubel, *et al.*, *"Current Protocols in Molecular Biology"*, Wiley & Sons, 1994.

Screening procedures which rely on nucleic acid hybridization make it possible to isolate any gene sequence from any organism, provided the appropriate probe is available. Oligonucleotide probes, which correspond to a part of the sequence encoding the protein in question, can be synthesized chemically. This requires that short, oligopeptide

stretches of amino acid sequence must be known. The DNA sequence encoding the protein can be deduced from the genetic code, however, the degeneracy of the code must be taken into account. It is possible to perform a mixed addition reaction when the sequence is degenerate. This includes a heterogeneous mixture of denatured double-stranded DNA. For such screening, hybridization is preferably performed on either single-stranded DNA or denatured double-stranded DNA.

Hybridization may be employed to screen recombinant clones by using labeled mixed synthetic oligonucleotide probes where each probe is potentially the complete complement of a specific DNA sequence in the hybridization sample which includes a heterogeneous mixture of denatured double-stranded DNA. For such screening, hybridization is preferably performed on either single-stranded DNA or denatured double-stranded DNA. Hybridization is particularly useful in the detection of cDNA clones derived from sources where an extremely low amount of mRNA sequences relating to the polypeptide of interest are present (cDNA is formed as a double-stranded DNA complement of mRNA). In other words, by using stringent hybridization conditions directed to avoid non-specific binding, it is possible, for example, to allow the autoradiographic visualization of a specific cDNA clone by the hybridization of the target DNA to that single probe in the mixture.

A cDNA library suspected of containing regions which encode a protein of interest can be screened by injecting the various mRNA derived from cDNAs into oocytes, allowing sufficient time for expression of the cDNA gene products to occur, and testing for the presence of the desired cDNA expression product, for example, by using probes for the motifs and a tissue expression pattern which is characteristic of Spp24 (see, e.g., SEQ.ID.Nos.1 through 6 and FIGURE 6 [tissue distribution of Spp24 in the bovine]). Alternatively, a cDNA library can be screened indirectly for peptides having at least one epitope in common with Spp24 using antibodies specific for the epitope of interest. As described below, such antibodies can be either polyclonally or monoclonally derived and used to detect expression product indicative of the presence of Spp24-homologous cDNA.

The probes of the present invention can also be used for examining the distribution of the specific fragments detected, as well as the quantitative (relative) degree of binding of the probe for determining the occurrence of specific strongly binding (hybridizing) sequences, thus indicating the likelihood for an individual to be at low risk or high risk for pathologies associated with excessive or inadequate inhibition of substrate by Spp24 or homologues thereof.

To that end, purified nucleic acid fragments containing intervening sequences or oligonucleotide sequences of 10-50 base pairs are radioactively labelled. The labelled preparations are used to probe nucleic acid from a histologic specimen by the Southern hybridization technique. Nucleotide fragments from a histologic specimen, before or after amplification, are separated into fragments of different molecular masses by gel electrophoresis and transferred to filters that bind nucleic acid. After exposure to the labelled probe, which will hybridize to nucleotide fragments containing target nucleic acid sequences, binding of the radioactive probe to target nucleic acid fragments is identified by autoradiography (see *Genetic Engineering*, 1, ed. Robert Williamson, Academic Press, (1981), 72-81).

Alternatively, nucleic acid from the specimen can be bound directly to filters to which the radioactive probe selectively attaches by binding nucleic acids having the sequence of interest. Specific sequences and the degree of binding is quantitated by directly counting the radioactive emissions.

Where the target nucleic acid is not amplified, detection using an appropriate hybridization probe may be performed directly on the separated mammalian nucleic acid. In those instances where the target nucleic acid is amplified, detection with the appropriate hybridization probe would be performed after amplification.

A probe may be detectably labelled with an atom or inorganic radical, most commonly using radionuclides, but also heavy metals can be used. Conveniently, a radioactive label may be employed. Radioactive labels include ^{32}P, ^{125}I, ^{3}H, ^{14}C, ^{111}In, ^{99m}Tc, or the like. Any radioactive label may be employed which provides for an adequate signal and has sufficient half-life. Other labels include ligands, which can serve as a specific binding pair member for a labelled ligand, and the like. A wide variety of labels routinely employed in immunoassays can readily be employed in the present assay. The choice of the label will be governed by the effect of the label on the rate of hybridization and binding of the probe to mutant nucleotide sequence. It will be necessary that the label provide sufficient sensitivity to detect the amount of mutant nucleotide sequence available for hybridization. Other considerations will be ease of synthesis of the probe, readily available instrumentation, ability to automate, convenience, and the like.

The manner in which the label is bound to the probe will vary depending upon the nature of the label. For a radioactive label, a wide variety of techniques can be employed. Commonly employed is nick translation with an a ^{32}P-dNTP or terminal phosphate hydrolysis with alkaline phosphatase followed by labeling with radioactive ^{32}P employing ^{32}P-NTP and T4 polynucleotide kinase. Alternatively, nucleotides can be synthesized where one or more of the elements present are replaced with a radioactive isotope, e.g., hydrogen with tritium. If desired, complementary labelled strands can be used as probes to enhance the concentration of hybridized label.

Where other radionucleotide labels are involved, various linking groups can be employed. For example, a terminal hydroxyl can be esterified, with inorganic acids, e.g., ^{32}P phosphate, or ^{14}C organic acids, or else esterified to provide linking groups to the label. Alternatively, intermediate bases may be substituted with activatable linking groups that can then be linked to a label.

Enzymes suitable for use as reporter groups will primarily be hydrolases, particularly esterases and glycosidases, or oxidoreductases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, and so forth. Chemiluminescers include luciferin, and 2, 3-dihydrophthalazinediones (e.g., luminol).

The probe can be employed for hybridizing to a nucleotide sequence affixed to a water insoluble porous support. Depending upon the source of the nucleic acid, the manner in which the nucleic acid is affixed to the support may vary. A representative hybridization protocol performed on nitrocellulose is described below; however, those of ordinary skill in the art will know of, or can easily identify, different supports that can be used in this respect. For example, hybridization techniques other than the representative example provided below are described by Gall and Pardue, (*Proc. Natl. Acad. Sci.* 63:378, 1969); and John, *et al.*, (*Nature*, 223:582, 1969).

Briefly, in a typical hybridization protocol, a nucleic acid from a histologic specimen is cloned and then spotted or spread onto a filter to provide a plurality of individual portions (plaques). The filter is an inert porous solid support, e.g., nitrocellulose. Any cells (or phage) present in the specimen are treated to liberate their nucleic acid. The lysing and denaturation of nucleic acid, as well as the subsequent washings, can be achieved with an appropriate solution for a sufficient time to lyse the cells and denature the nucleic acid. For lysing, chemical lysing will conveniently be employed, as described previously for the lysis buffer. Other denaturation agents include elevated temperatures, organic reagents, e.g., alcohols, amides, amines, ureas, phenols and sulfoxides or certain inorganic ions, e.g., thiocyanate and perchlorate.

After denaturation, the filter is washed in an aqueous buffered solution, such as Tris, generally at a pH of about 6 to 8, usually 7. One or more washings may be involved, conveniently using the same procedure as employed for the lysing

and denaturation. After the lysing, denaturing, and washes have been accomplished, the nucleic acid spotted filter is dried at an elevated temperature, generally from about 50°C to 70°C. Under this procedure, the nucleic acid is fixed in position and can be assayed with the probe when convenient.

Pre-hybridization may be accomplished by incubating the filter with the hybridization solution without the probe at a mildly elevated temperature for a sufficient time to thoroughly wet the filter. Various hybridization solutions may be employed, comprising from about 20% to 60% volume, preferably 30%, of an inert polar organic solvent. A common hybridization solution employs about 50% formamide, about 0.5 to 1M sodium chloride, about 0.05 to 0.1M sodium citrate, about 0.05 to 0.2% sodium dodecylsulfate, and minor amounts of EDTA, ficoll (about 300-500 kD), polyvinylpyrrolidone, (about 250-500 kD) and serum albumin. Also included in the hybridization solution will generally be from about 0.5 to 5 mg/ml of sonicated denatured DNA, e.g., calf thymus of salmon sperm; and optionally from about 0.5 to 2% wt/vol glycine. Other additives may also be included, such as dextran sulfate of from about 100 to 1,000 kD and in an amount of from about 8 to 15 weight percent of the hybridization solution.

The amount of labelled probe present in the hybridization solution will vary widely, depending upon the nature of the label, the amount of the labelled probe that can reasonably bind to the filter, and the stringency of the hybridization. Generally, substantial excess over stoichiometric concentrations of the probe will be employed to enhance the rate of binding of the probe to the fixed target nucleic acid.

Various degrees of stringency of hybridization may be employed. The more severe the conditions, the greater the complementarily that is required for hybridization between the probe and the single stranded target nucleic acid sequence for duplex formation. Severity can be controlled by temperature, probe concentration, probe length, ionic strength, time, and the like.

Conveniently, the stringency of hybridization is varied by changing the polarity of the reactant solution by manipulating the concentration of formamide in the range of 20% to 50%. Temperatures employed will normally be in the range of about 20°C to 80°C, usually 30°C to 75°C .

After the filter has been contacted with a hybridization solution at a moderate temperature for a period of time sufficient to allow hybridization to occur, the filter is then introduced into a second solution having analogous concentrations of sodium chloride, sodium citrate and sodium dodecylsulfate as provided in the hybridization solution. The time the filter is maintained in the second solution may vary from five minutes to three hours or more. The second solution determines the stringency, dissolving cross duplexes and short complementary sequences. After rinsing the filter at room temperature with dilute sodium citrate-sodium chloride solution, the filter may now be assayed for the presence of duplexes in accordance with the nature of the label. Where the label is radioactive, the filter is dried and exposed to X-ray film.

The label may also comprise a fluorescent moiety that can then be probed with a specific antifluorescent antibody. For example, horseradish peroxidase enzyme can be conjugated to this antibody to catalyze a chemiluminescent reaction. Production of light can then be seen on rapid exposure to film.

Based on the degree of sequence homology identified between Spp24 and cystatin domain proteins obtained from non-bovine and non-human species, it can be reasonably predicted that Spp24 homologs can be identified in cells from other mammalian species by, for example, utilization of the techniques summarized above.

B. Expression of Spp24 Encoding Polynucleotides.

In the present invention, the polynucleotides encoding Spp24, or polynucleotides which operatively encode fragments thereof, may be inserted into a recombinant expression vector. The term "recombinant expression vector" refers to a plasmid, virus or other vehicle known in the art that has been manipulated by insertion or incorporation of the appropriate genetic sequences. Such expression vectors contain a promoter sequence which facilitates the efficient transcription of the inserted genetic sequence of the host; a polynucleotide which will encode a peptide under the control of such a promoter is an "operatively encoding" polynucleotide.

Transformation of a host cell with recombinant DNA may also be carried out by conventional techniques as are well known to those skilled in the art. Where the host is prokaryotic, such as *E. coli*, competent cells which are capable of DNA uptake can be prepared from cells harvested after exponential growth phase and subsequently treated by the $CaCl_2$ method by procedures well known in the art. Alternatively, $MgCl_2$ or RbCl can be used. Transformation can also be performed after forming a protoplasm to the host cell or by electroporation.

For review regarding conventional techniques for development and use of recombinant expression vectors, those of skill in the art may wish to refer to Ausubel, *et al., "Current Protocols in Molecular Biology"*, Wiley & Sons, 1994.

For certain applications, it may be desirable to develop modified forms of Spp24 by mutation of polynucleotides which operatively encode Spp24 or a fragment thereof. Such modified gene products may be generated to stabilize the expression product, optimize the pH requirements of the expression product, or reduce its immunogenecity. Most such modifications involve single base substitutions of nucleotides not required for the biological activity of the expression product. Modifications of polynucleotide sequences may also be

employed to generate a set of fully degenerate probes for use in hybridization protocols, as discussed above.

For example, a polynucleotide having a desirable mutation may be produced by site-specific mutagenesis using a conventional polymerase chain reaction (PCR) and a primer pair corresponding to the 3' and 5' regions of the cDNA. A preferred method of mutation-generating PCR amplification is the overlap extension PCR technique described by Ho, *et al., Gene* 77:51-59 (1989), the disclosure of which is incorporated herein by this reference. Generally, this technique accomplishes site-specific mutagenesis of the clone by utilizing a 3' primer to add the mismatched mutating bases (primer B in the Ho article, which is used with the 5' primer A in the first PCR cycle described). Amplification using the A and B primers yields an AB fragment. A second PCR cycle uses a primer (D) from the 3' end of the gene and a 5' mutated primer (C) complementary to primer B. The resulting amplification product (fragment CD) will overlap the AB fragment. When the AB and CD fragments are denatured, reannealed and amplified using the A and D primers, the resulting fusion product (AD) will contain the full-length cDNA sequence and the desired mutation.

For generation of random mutations in genes, a PCR method modified to decrease the fidelity of Taq polymerase during DNA synthesis is particularly useful. Specifically, the PCR is performed under conditions that will generally produce a cumulative error frequency of 2% over the entire DNA nucleotide sequence and a mutant yield (for targets over 300 bp) of greater than 90%; to wit, conditions of high and pool-biased dNTP concentrations (1mM each dTTP, dCTP, dGTP, 200 μM dATP), a high concentration of $MgCl_2$ in the presence of $MnCl_2$ (0.5mM), and 25 PCR cycles starting with 1ng cloned plasmid target. For a further review of this protocol for randomized PCR mutagenesis, those of skill in the art may wish to consult Leung, *et al., Technique,* 1:11-15, 1989; also, for a modification of the method to achieve random point mutations, Cadwell and Joyce, *PCR Methods and Applications*, 2:28-33, 1992; and, for generation of a

PCR-derived library of random point mutations, Kirchoff and Desrosiers, *PCR Methods and Applications,* 2:301-304, 1993, the disclosures of which are incorporated herein by this reference to illustrate PCR mutagenesis methods known in the art.

Another suitable approach to making single base substitutions or deletions is described by Shaw in U.S. Patent No. 4,904,584 ("Site-Specific Homogeneous Modification of Polypeptides"), the disclosure of which is incorporated herein by this reference for purposes of illustrating knowledge in the art regarding methods for achieving specific mutations in a nucleic acid sequence.

C. Spp24 Protein and Peptide Fragments Thereof.

As discussed in the Background section above, there is a substantial degree of homology among the cystatin domain peptides identified in FIGURE 7 and Spp24. It can therefore be reasonably predicted that the biological function of Spp24 is essentially the same as the peptides identified in FIGURE 7; i.e., a cysteine protease inhibitor (see also, discussion in Example V, below). Further, based on the tissue distribution of Spp24 in bovine tissues, it can also be reasonably predicted that Spp24 is active in the metabolism of mammalian tissue, particularly bone (see, FIGURE 6 and the discussions in Examples III and V, below). Thus, as further detailed below, Spp24 can be reasonably predicted to have value in screening for metabolic (particularly osteoinductive) protease activity and may have therapeutic value in controlling abnormal bone metabolism.

Compared to isolation and purification of native Spp24 or homologs thereof from mammalian cells, production of recombinant Spp24 or fragments thereof will ensure greater availability of the desired peptides. To that end, isolation and purification of microbially expressed Spp24, or fragments thereof, may be carried out by conventional means including preparative chromatography and

immunological separations involving monoclonal or polyclonal antibodies. "Isolated, substantially pure Spp24" will therefore mean Spp24, whether native, recombinant or synthetic, which is substantially free of components and contaminants which may be associated with the native protein.

Alternatively, Spp24 and fragments thereof may be duplicated synthetically. To that end, based on the information contained in SEQ. ID. No. 7, the deduced full-length amino acid sequence for Spp24 may be readily duplicated. Using this information, Spp24 peptides may also be synthesized without undue experimentation by commonly used methods such as t-BOC or FMOC protection of alpha-amino groups. Both methods involve stepwise synthesis whereby a single amino acid is added at each step starting from the C terminus of the peptide (see, Coligan, *et al., Current Protocols in Immunology,* Wiley Interscience, 991, Unit 9). Peptides of the invention can also be synthesized by various well known solid phase peptide synthesis methods, such as those described by Merrifield, *J. Am. Chem. Soc.*, 85:2149 (1962), and Stewart and Young, *Solid Phase Peptides Synthesis*, (Freeman, San Francisco, 27-62, 1969), using a copoly(styrene-divinylbenzene) containing 0.1-1.0 mMol amines/g polymer.

In this latter method, completion of chemical synthesis, the peptides can be deprotected and cleaved from the polymer by treatment with liquid HF-10% anisole for about 1/4-1 hours at 0°C. After evaporation of the reagents, the peptides are extracted from the polymer with 1% acetic acid solution which is then lyophilized to yield the crude material. This can normally be purified by such techniques as gel filtration on Sephadex G-15 using 5% acetic acid as a solvent. Lyophilization of appropriate fractions of the column will yield the homogeneous peptide or peptide derivatives, which can then be characterized by such standard techniques as amino acid analysis, thin layer chromatography, high performance liquid chromatography, ultraviolet absorption spectroscopy, molar rotation, solubility, and quantitated by the solid phase Edman degradation.

Peptides of the invention include functional derivatives of Spp24 and fragments thereof. By "functional derivative" is meant the "fragments," "variants," "analogs," or "chemical derivatives" of a molecule. A "fragment" of a molecule, includes any amino acid subset of the molecule. A "variant" of such molecule refers to a naturally occurring molecule substantially similar to either the entire molecule, or a fragment thereof. An "analog" of a molecule refers to a non-natural molecule substantially similar to either the entire molecule or a fragment thereof.

A molecule is said to be "substantially similar" to another molecule if the sequence of amino acids in both molecules is substantially the same. Substantially similar amino acid molecules will possess a similar biological activity. Thus, provided that two molecules possess a similar activity, they are considered variants as that term is used herein even if one of the molecules contains additional amino acid residues not found in the other, or if the sequence of amino acid residues is not identical.

As used herein, a molecule is said to be a "chemical derivative" of another molecule when it contains additional chemical moieties not normally a part of the molecule. Such moieties may improve the molecule's solubility, absorption, biological half life, etc. The moieties may alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, etc. Moieties capable of mediating such effects are disclosed, for example, in *Remington's Pharmaceutical Sciences*, 16th Ed., Mack Publishing Co., Easton, Penn. (1980).

Minor modifications of the Spp24 primary amino acid sequence may result in proteins which have substantially equivalent activity as compared to the Spp24 peptides described herein. For example, deletion of one or more amino acids can also result in a modification of the structure of the resultant molecule without significantly altering its biological activity. This can lead to the development of

a smaller active molecule which would have broader utility. Also, one can remove amino or carboxy terminal amino acids which may not be required for the protein to exert the desired biologic or antigenic activity.

The term "conservative variation" as used herein denotes the replacement of an amino acid residue by another, biologically similar residue. Examples of conservative variations include the substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as the substitution of arginine for lysine, glutamic for aspartic acids, or glutamine for asparagine, and the like. The term "conservative variation" also includes the use of a substituted amino acid in place of an unsubstituted parent amino acid provided that the antibodies raised to the substituted polypeptide also immunoreact with the unsubstituted polypeptide.

Such variations in amino acid sequence may also be produced in high volume through expression of recombinant expression vectors containing polynucleotides modified to encode such variations. Such modifications may be deliberate, as by site-directed mutagenesis of an operatively encoding polypeptide, or may be spontaneous. Methods for accomplishing such mutations are discussed above. All of the peptides produced by these modifications are included herein as long as the biological activity of Spp24 still exists; e.g., Spp24 will bind a complementary binding molecule. In this respect, " binding molecule" refers to a molecule which specifically binds Spp24 and may include antibodies as well as substrate molecules (i.e., protease). Such molecules may be identified by assay protocols which are well known to those skilled in the art; examples thereof are discussed in greater detail below.

D. Production and Use of Anti-Spp24 Antibodies.

Antibodies to Spp24, fragments thereof, and/or homologs thereof for use in detecting Sp24 or homologs of Spp24 in a sample, as well as for therapeutic use, can be produced as described below.

The antigenicity of Spp24 peptides can be determined by conventional techniques to determine the magnitude of the antibody response of an animal which has been immunized with the peptide. Generally, the Spp24 peptides which are used to raise the anti-Spp24 antibodies should generally be those which induce production of high titers of antibody with relatively high affinity for Spp24.

Once antigenic peptides are prepared, antibodies to the immunizing peptide are produced by introducing peptide into a mammal (such as a rabbit, mouse or rat). A multiple injection immunization protocol is preferred for use in immunizing animals with the antigenic Spp24 peptides (see, e.g., Langone, *et al.*, eds., "Production of Antisera with Small Doses of Immunogen: Multiple Intradermal Injections", *Methods of Enzymology* (Acad. Press, 1981). For example, a good antibody response can be obtained in rabbits by intradermal injection of 1 mg of an antigenic Spp24 peptide emulsified in Complete Freund's Adjuvant followed several weeks later by one or more boosts of the same antigen in Incomplete Freund's Adjuvant.

If desired, the immunizing peptide may be coupled to a carrier protein by conjugation using techniques which are well-known in the art. Such commonly used carriers which are chemically coupled to the peptide include keyhole limpet hemocyanin (KLH), thyroglobulin, bovine serum albumin (BSA), and tetanus toxoid. The coupled peptide is then used to immunize the animal (e.g. a mouse or a rabbit). Because Spp24 is presently believed to be highly conserved

among vertebrate species, use of a carrier protein to enhance the immunogenecity of Spp24 proteins is preferred.

Polyclonal antibodies produced by the animals can be further purified, for example, by binding to and elution from a matrix to which the peptide to which the antibodies were raised is bound. Those of skill in the art will know of various techniques common in the immunology arts for purification and/or concentration of polyclonal antibodies, as well as monoclonal antibodies (see, for example, Coligan, *et al.*, Unit 9, *Current Protocols in Immunology,* Wiley Interscience, 1991).

For their specificity and ease of production, monoclonal antibodies are preferred for use in detecting Spp24 in a sample. For preparation of monoclonal antibodies, immunization of a mouse or rat is preferred. The term "antibody" as used in this invention is meant also to include intact molecules as well as fragments thereof, such as for example, Fab and $F(ab')_2$, which are capable of binding the epitopic determinant. Also, in this context, the term "mAb's of the invention" refers to monoclonal antibodies with specificity for Spp24.

The general method used for production of hybridomas secreting monoclonal antibodies ("mAb's"), is well known (Kohler and Milstein, *Nature*, 256:495, 1975). Briefly, as described by Kohler and Milstein the technique comprised lymphocytes isolated from regional draining lymph nodes of five separate cancer patients with either melanoma, teratocarcinoma or cancer of the cervix, glioma or lung, were obtained from surgical specimens, pooled, and then fused with SHFP-1. Hybridomas were screened for production of antibody which bound to cancer cell lines.

Confirmation of Spp24 specificity among mAb's can be accomplished using relatively routine screening techniques (such as the enzyme-linked immunosorb-

ent assay, or "ELISA") to determine the elementary reaction pattern of the mAb of interest.

It is also possible to evaluate an mAb to determine whether it has the same specificity as a mAb of the invention without undue experimentation by determining whether the mAb being tested prevents a mAb of the invention from binding to Spp24 isolated as described above. If the mAb being tested competes with the mAb of the invention, as shown by a decrease in binding by the mAb of the invention, then it is likely that the two monoclonal antibodies bind to the same or a closely related epitope.

Still another way to determine whether a mAb has the specificity of a mAb of the invention is to pre-incubate the mAb of the invention with an antigen with which it is normally reactive, and determine if the mAb being tested is inhibited in its ability to bind the antigen. If the mAb being tested is inhibited then, in all likelihood, it has the same, or a closely related, epitopic specificity as the mAb of the invention.

Spp24, Spp24 homologs, and/or antibodies thereto of the invention are particularly suited for use in immunoassays in which they can be utilized in liquid phase or bound to a solid phase carrier. In addition, such compounds can be detectably labelled in a variety of ways which are well-known in the art. Immunoassays which utilize such compounds can be applied to identify Spp24 in a biological sample, to purify Spp24 from a biological sample, to screen for binding molecules which will bind to Spp24, and to determine the extent of Spp24 binding molecule activity in a biological sample (i.e., a sample of mammalian tissue or body fluid, such as bone tissue, synovial fluid or serum).

Examples of such immunoassays are competitive and non-competitive immunoassays in either a direct or indirect format. More specifically such immunoassays include the radioimmunoassay (RIA), the sandwich

(immunometric assay) and the Western blot assay. Detection of antibodies which bind to Spp-24 or Spp-24 homologs can be done utilizing immunoassays which run in either the forward, reverse, or simultaneous modes, including immunohistochemical assays on physiological samples. The concentration of materials used in the assay will vary depending on the type of immunoassay and nature of the detectable label which is used and can be readily determined by one of ordinary skill in the art using routine experimentation.

Examples of well-known carriers for use in such immunoassays include glass, polystyrene, polyvinyl chloride, polypropylene, polyethylene, polycarbonate, dextran, nylon, amyloses, natural and modified celluloses, polyacrylamides, agaroses, and magnetite. The nature of the carrier can be either soluble or insoluble for purposes of the invention. Those skilled in the art will know of other suitable carriers for binding Spp24 or Spp24 homologs, or will be able to ascertain such, using routine experimentation.

Further, there are many different labels and methods of labeling known to those of ordinary skill in the art which can be utilized in such immunoassays. Examples of the types of labels which can be used in the present invention include enzymes, radioisotopes, colloidal metals, fluorescent compounds, chemiluminescent compounds, and bioluminescent compounds. Methods for binding such labels to peptides and/or antibodies will be well-known to, or can be readily determined by, those of ordinary skill in the art.

Potential Spp24 binding molecule molecules may also be screened and identified using substantially pure Spp24 of the invention. In general, Spp24 will be incubated with candidate binding molecules and the interaction (or lack thereof) between the molecules evaluated by determining whether the activity of the candidate binding molecule is altered in the presence of Spp24 (or peptide fragments thereof). Those of skill in the art will recognize, however, that

binding molecule binding to Spp24 may also be characterized by determination of other parameters, such as binding kinetics and affinity.

Once a particular binding molecule has been determined to bind Spp24, other binding molecules may be screened for binding by inhibition and/or competitive binding studies as described above with respect to screening of mAb's with specificity for Spp24.

E. Spp24 Compositions and Methods for *In Vivo* Use Thereof.

As indicated in the Background discussion herein, bone resorption is implicated in several known pathological states, including osteoarthritis. Control of bone resorption may therefore be enhanced by including Spp24 (to inhibit proteolysis to a greater degree than achieved in the host prior to treatment), or anti-Spp24 antibodies (for *in vivo* absorption of Spp-24 to allow a greater degree of proteolysis to occur in the host than occurred prior to treatment), in a pharmaceutically acceptable composition to be administered to the host or, for localized repair of injuries to bone, to be contacted with the site of injury (by, for example, incorporation into an osteoinductive carrier).

To that end, Spp-24, Spp-24 homologs, or anti-Spp-24 antibodies of the invention may be formulated in a pharmaceutically carrier. Such carriers include sterile aqueous of non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media.

Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those

based on Ringer's dextrose), and the like. Preservatives and other additives may also be present such as, for example, antimicrobials, antioxidants, chelating agents, and inert gases and the like.

Exemplary of solid carriers are lactose, terra alba, sucrose, talc, gelatin, agar, pectin, acacia, magnesium stearate, stearic acid, microcrystalline cellulose, polymer hydrogels and the like. Similarly, the carrier or diluent may include any time delay material well known to the art, such as glyceryl monostearate or glyceryl distearate alone or with a wax, microcapsules, microspheres, liposomes, and hydrogels. For further reference, those of skill in the art may wish to consult the standard reference *Remington's Pharmaceutical Sciences* (which is incorporated herein by reference to illustrate knowledge in the art concerning suitable pharmacuetical carriers).

A wide variety of pharmaceutical forms can be employed. Thus, when using a solid carrier the preparation can be tableted (however, the oral route of administration should be avoided due to first pass metabolic degradation), placed in a hard gelatin capsule in powder or pellet form, or in the form of a troche, lozenge or suppository. When using a liquid carrier the preparation can be in the form of a liquid, such as an ampule, or as an aqueous or nonaqueous liquid suspension. Topical administration via timed release skin patches is also a suitable pharmaceutical form.

The efficacy of such treatments will preferably be monitored by tracking levels of Spp24, Spp24 binding molecule and bone resporption in the host, as well as by monitoring other clinical signs of health well known to those of ordinary skill in the clinical art.

Dosages of Spp24, Spp24 homologs, and/or Spp24 antibodies for *in vivo* administration will vary depending on the age, weight and presenting condition of the host to be treated. Such variables will readily be accounted for by those

of ordinary skill in the clinical art. In particular, dosages will be adjusted upward or downward for each recipient based on the severity of the condition to be treated and accessibility of the target cells to the pharmaceutical formulations of the invention. Where possible, it will be preferable to administer the pharmaceutical formulations of the invention locally at the site of the target cells; e.g., into inflamed skin or by infusion to another organ of the host. Thus, dosages will also vary depending on the route of administration and the extent to which the formulations of the invention are expected to reach target cells before dilution or clearance of the formulation.

Preferred routes of administration of the therapeutic compositions of the invention are by local injection or parenteral infusion, although oral and intravascular routes may also be utilized. Such means will be particularly suited for treatment of pathologies associated with non-localized excessive or inadequate inhibition of cysteine proteolysis, such as osteoarthritis at more than one discrete body region in the host, or loss of connective tissue in the liver.

Generally, however, based on experience with other inhibitors of cysteine proteases and the data provided herein, good results can be expected to be achieved in an adult host of about 60 kg. body weight in a dosage range of about 500 to about 4,000 mg/day, preferably between about 1,000 and about 3,500 mg/day (i.e., a "therapeutically effective dosage"). These dosages may be combined with other conventional pharmaceutical therapies for the condition to be treated in the host.

For use in repair of localized tissue injury (particularly in bone), Spp-24 and/or Spp24 homologs may be conveniently administered by incorporation into a osteosupportive scaffold. Such scaffolds are known in the art; examples thereof are described in U.S. Patent Nos. 5,041,188 and 4,186,443, the disclosures of which are incorporated herein by this reference to illustrate the state of the art

concerning the use of such scaffolds and incorporation of osteoregulatory factors therein.

The invention having been fully described, examples illustrating its practice are set forth below. These examples should not, however, be considered to limit the scope of the invention, which is defined by the appended claims.

In the examples, the abbreviation "min." refers to minutes, "hrs" and "h" refer to hours, and measurement units (such as "ml") are referred to by standard abbreviations.

EXAMPLE I
PURIFICATION AND CHARACTERIZATION OF BOVINE SPP24

Non-collagenous proteins were first extracted from 300 g of ground calf bone by demineralization with 3L of 10% formic acid for 3 h at 4°C. The extracted proteins were then absorbed to a C_{18} matrix, washed with a 0.1% trifluoroacetic acid to remove bone mineral, and eluted with 50% acetonitrile in 0.1% trifluoroacetic acid. After freeze drying, the neutral pH-soluble proteins were removed by repeated suspension in 25 ml of 50 mM NH_4HCO_3 followed by centrifugation until the supernatant A_{280} fell below 0.1.

The neutral pH-insoluble pellet was dissolved in 3 ml of 6 M guanidine Hcl in 0.1 M Tris pH 9 and applied to a 2 x 150 cm column of Sephacryl S-100 HR equilibrated with the same buffer at room temperature. The eluant fractions containing Spp24 were pooled and further purified using a 4.6mm x 25cm C_4 reverse phase HPLC column with a 2 h gradient from 0.1% trifluoroacetic acid in water to 0.1% trifluoroacetic acid in 60% acetonitrile at a flow rate of 1 ml/min.

SDS-polyacrylamide gel electrophoresis to identify by Spp24 was performed under reducing conditions using 4-20% gradient gels (Novex, San Diego, CA).

Spp24 insoluble at neutral pH and can be separated from neutral pH-soluble proteins such as the bone Gla protein by repeated washing of the dried proteins in the acid extract with 50 mM NH_4HCO_3. The neutral pH-insoluble protein fraction can be solubilized in 6 M guanidine HCL buffer and fractionated by gel filtration over Sephacryl S-100 HR™. As can be seen in FIGURE 2, the major protein component recovered from the gel filtration step is MGP. The MGP isolated by this procedure is pure by the criteria of SDS gel electrophoresis in 18% polyacrylamide gels and by N-terminal protein sequencing (data not shown). Spp24 is recovered in a single peak centered at fraction 62 and, as

judged by SDS gel electrophoresis, is somewhat contaminated with other proteins (FIGURE 2 inset).

These protein contaminants can be removed from Spp24 by chromatography on a C4 HPLC column using a linear acetonitrile gradient (FIGURE 3). The resulting purified Spp24 is homogeneous by the criteria of SDS gel electrophoresis (FIGURE 3, inset) and N-terminal protein sequencing. To obtain additional sequence data for the construction of degenerate primers for RT-PCR, purified Spp24 was cleaved at methionine with cyanogen bromide and the resulting peptides were purified by filtration over Sephacryl S-100 HR in 6 M guanidine HCL buffer and subjected to N-terminal protein sequencing. The two internal peptides isolated by these procedures yielded single N-terminal sequences.

The coding region of Spp24 is terminated by a single TGA triplet at nucleotides 691-693 and a TAA triplet follows at nucleotides 700-702. The 3'-untranslated region consists of 139 nucleotides with a polyadenylation signal (AATAAA) at nucleotides 790-795. The ATG, found at nucleotides 91-93, was considered to be the initiation codon according to conventional rules for translation initiations. The open reading frame codes for a 200-residue long protein containing a 20-residue transmembrane signal peptide with a potential signal peptidase cleavage site at amino acid residue 20.

The N-terminus of mature Spp24 was identified by N-terminal sequencing (see, SEQ. ID NO:7) and is located 20 amino acids downstream from the presumed initiation methionine. The deduced mature form of Spp24 contains a total of 180 amino acids and has a calculated molecular weight of 20,458, in agreement with the 24 kDa size of Spp24 determined by SDS-PAGE (FIGURES 2 and 3; SEQ. ID. NO: 7).

Purified proteins and peptides were transferred to PVDF membranes using a ProSpin™ device (Applied Biosystems, Foster City, CA) and sequenced using an Applied Biosystems 470 A sequenator equipped with a model 120 on-line HPLC. In order to determine the location of phosphoserine residues were converted to S-ethylcysteine by reaction with ethanethiol. The percent phosphorylation at each target serine residue was determined from the relative recovery of serine and S-ethylcysteine at that sequence position (see, N-terminus of SEQ. ID. NO:7).

EXAMPLE II

DETERMINATION OF cDNA NUCLEOTIDE SEQUENCE ENCODING BOVINE Spp24

Reverse transcription of RNA and polymerase chain reaction amplification of cDNA (RT-PCR) was used to determine the cDNA sequence of Spp24. Based on information from the N-terminal amino acid sequence of purified bovine bone Spp24 and of the cyanogen bromide peptide 1, a N-terminal sense degenerate primer (5'-TTT/CCCIGTITAT/CGAT/CTAT/CGA-3', where I stands for inosine; SEQ. ID. NO:1) and an internal antisense degenerate primer (5'-AA/GIATA/GTCICCA/GAAA/GAACAT-3'; SEQ. ID NO:2) were designated. Using an RNA preparation from either bovine bone periosteum (plus adjacent bone scrapings) or bovine liver and a GeneAmp RNA PCR Kit (PERKIN ELMER CETUS, Norwalk, CT), a 380-bp cDNA fragment was generated and then amplified by RT-PCR in a Perkin-Elmer DNA thermal cycler. The PCR product was cloned directly into a plasmid vector using the TA Cloning™ System (INVITROGEN, San Diego, CA). Following the enzymatic dideoxy chain termination method, both strands of the plasmid cDNA inserts were sequenced with a Version 2.0 DNA Sequencing Kit, (United States Biochemical, Cleveland, OH) using a 5% Long Ranger gel (AT Biochem, Malvern, PA). Identical sequences were obtained for this 380-bp fragment from liver and bone. The plasmid cDNA inserts were excised with EcoRI. Due to an internal EcoRI site

in the insert, this digestion produced 68-bp and 312-bp cDNA fragments. The 312-bp fragment was employed as a probe for screening of a bovine liver λgt11 cDNA library and for Northern blots.

The message from the internal region to the 3'-end was sequenced after generating a PCR fragment using 3'-Rapid Amplification of cDNA Ends (3'-RACE). A specific internal sense primer (5'-CGCTGCCACTGGTCCTCCAGCTCT-3'; SEQ. ID. NO:3) was synthesized which was located 45-bp upstream from the internal antisense degenerate primer. In addition, a unique 23-base oligonucleotide adapter primer linked to a 17-oligo(dt) (5'-ACGCGTCGACCTCGAGATCGATG-$(dT)_{17}$ -3'; SEQ. ID. NO:4), and the adapter primer (5'-ACGCGTCGACCTCGAGATCGATG-3'; SEQ. ID. NO:5) were used for the 3'-RACE system. A 370-bp cDNA fragment was produced by both bovine periosteum and bovine liver total RNA and subsequently cloned and sequenced as described above. Identical sequences were obtained for this 370-bp fragment from bone and liver.

The 5'-end cDNA clone was obtained by screening a bovine liver λgt11 cDNA library (CLONTECH, Palo Alto, CA) with the 312-bp cDNA probe. Bacteriophages were plated and transferred to nitrocellulose filters as described by the Clontech protocol. Two replica filters were lifted from each plate. Following hybridization to the ^{32}P-labeled bovine 312-bp cDNA probe, eight positive phage plaques were isolated. Two of these contained the 5'-end 380-bp cDNA sequence as determined by PCR with the two degenerate primers. cDNA inserts of these two clones were enzymatically amplified by performing PCR with a specific antisense primer (5'-CAGATAGGGGCTCATGACTGGGA-3'; SEQ ID NO:6) located 73-bp downstream from the N-terminal sense degenerate primer and either a λg11 forward or a λgt11 reverse primer (PROMEGA, Madison, WI). A 470-bp PCR product was obtained from one clone and a 300-bp PCR product from the other clone. The 5'-end PCR products were cloned

and partially sequenced with TA Cloning™ System and Version 2.0 DNA Sequencing Kit as above (see, e.g., FIGURE 4).

The complete cDNA nucleotide sequence for Spp24 determined as above is shown in SEQ. ID NO:7.

EXAMPLE III
TISSUE DISTRIBUTION OF Spp24 IN BOVINE TISSUES

Bovine steer ankle bones were obtained from Talone's Meat Packers (Escondido, CA). The periosteum and adjacent bone scrapings were removed and frozen in liquid nitrogen. Frozen bovine heart, lung, kidney, liver and spleen were purchased from Monfort Biological (Greeley, CO).

Total RNA was isolated from the tissues using guanidine isothiocyanate and fractionated on 1.4% agarose formaldehyde gels in a 3-[N-morpholino] propane sulfonic acid buffer and transferred to a nylon membrane (0.45 μm Nytran™; Schleicher and Schuell, Keene, NH). The 312-bp cDNA probe was labeled with [α-^{32}P]dCTP using a Random Primed DNA Labeling Kit (Boehringer Mannheim, Indianapolis, IN) to a specific activity of at least 1 x 10^8 cpm/μg. Hybridization to RNA immobilized on Nytran was typically performed for 17h at 42°C in the same solution without the cDNA probe. Filters were washed in 0.5 x SSC (1x SSC is 180 mM NaCl and 15 mM Na Citrate) containing 0.2% sodium dodecyl sulfate for 10 min. at room temperature, followed by 1 h at 55°C. Autoradiography was performed at -70°C with Kodak XAR-5 film (Eastman Kodak, Rochester, NY). RNA size markers (GIBCO BRL, Gaithersburg, MD) were run on adjacent gel lanes and stained with methylene blue.

Spp24 was determined as above to be predominately present in bone and liver. A copy of the above-described blot is contained in FIGURE 6. The size of the Spp24 mRNA from bovine liver and bone (periosteum) agrees with the size of the Spp24 cDNA (SEQ. ID. NO: 7).

EXAMPLE IV
IDENTIFICATION OF PHOSPHOSERINE IN THE SERINE-RICH SEQUENCE OF BOVINE Spp24 PROTEIN

As can be seen in SEQ. ID NO:7, there is a serine-rich sequence between residues 128 and 136 of Spp24 which contains several potential sites of serine phosphorylation by the Ser-X-Glu/Ser(P)-specific secretory pathway protein kinase. In order to evaluate the possible phosphorylation of these serine residues, Spp24 was first cleaved at tryptophan 127 using BNPS-skatole and the peptide corresponding to residues 128-190 of Spp24 was purified by gel filtration over Sephacryl S-100 HR. This peptide was then treated with ethanethiol in order to convert the putative phosphoserine residues to S-ethylcysteine (14,18) and subjected to N-terminal protein sequencing. The results of this analysis unambiguously identified the PTH derivative of S-ethylcysteine at every serine residue with the serine rich region of Spp24. As has been noted for all other phosphoproteins secreted into the extracellular environment of cells, the extent of serine phosphorylation was in each instance partial, ranging from 5% to 83%. The presence of phosphoserine in Spp24 was confirmed by conventional techniques for acid hydrolysis and amino acid analysis.

EXAMPLE V
HOMOLOGY OF BOVINE Spp24 PROTEIN WITH OTHER CYSTEINE-PROTEASE INHIBITORS

To evaluate the possible relationships between Spp24 and other known proteins, the complete 200 residue Spp24 sequence deduced from its cDNA structure was compared with all presently known protein sequences in the non-redundant database of the NLM using the BLAST search program. This search revealed the presence of a comparable level of sequence identity between Spp24 and cystatin domains 1 and 3 of human kininogen and between Spp24 and the precursor to the bovine neutrophil antibiotic peptide bactenecin (FIGURES 1 and

7; SEQ.ID.No. 1). The bactenecin precursor and cystatin domains 1 and 3 of kininogen are known to be related in sequence to the cystatin family of thiol protease inhibitors, and Spp24 is accordingly compared to two additional members of this family in FIGURE 7, porcine cathelin and chicken cystatin.

As can be seen by analysis of FIGURE 7, the bactenecin precursor and cathelin are more closely related to Spp24 than to cystatin domains 1 and 3 of kininogen or to chicken cystatin. Cystatin domains 1 and 3 of kininogen and cystatin are also more closely related to Spp24 than to bactenecin precursor or to cathelin. It is therefore probable that Spp24 is an evolutionary intermediate which links cathelin, bactenecin precursor, and the closely related precursors to the neutrophil antibiotic peptides Bac5 and indolicidin with the various cystatins and with the cystatin domains of kininogen.

The structure of a prototypical cystatin domain has been determined from crystallographic studies of the 108 residue chicken cystatin, and is a compact structure with a 5 stranded β-sheet wrapped around a 5 turn α-helix. It is highly probable that the sequence identities observed between Spp24 and chicken cystatin reflect similar polypeptide conformations and that the entire 107 residue region of Spp24 between the N-terminus of the mature protein and the 11 residue phosphoserine-rich sequence is folded into a cystatin-like tertiary structure. Since the four cysteine residues in the cystatin domain of Spp24 lie at sequence positions known to be involved in disulfide bonds in other members of the cystatin family, it is also probable that these four cysteine residues are likewise involved in disulfide bonds, and that these bonds join Cys 63 with Cys 94 and Cys 87 with Cys 105 in the mature Spp24 protein (FIGURE 5; SEQ. ID. NO:7).

Among the proteins most closely related in sequence to Spp24, cathelin, chicken cystatin, and cystatin domain 3 of kininogen have been previously shown to inhibit thiol proteases (25,26). Although the bactenecin precursor has not itself

been tested for its ability to inhibit thiol proteases, the closely related precursor to the neutrophil antibiotic peptide Bac5 has been reported to potently inhibit cathepsin. Several thiol proteases of the cathepsin family are in fact known to be expressed by bone cells, and there is evidence to suggest that such thiol proteases may be released from osteoclasts in order to digest collagen and various non-collagenous proteins under the acidic conditions of osteoclast-mediated bone resorption. It is therefore highly probable that Spp24 is a thiol protease inhibitor.

This conclusion is further supported by the presence of several intriguing similarities between Spp24 and fetuin (FIGURE 7) which suggest that the proteins could have similar mechanisms of action in bone. Both proteins are synthesized by liver as well as bone, and accumulate in the extracellular matrix of bone. Both proteins have cystatin domains, one for Spp24 and two for fetuin. Both proteins contain phosphoserine. Finally, both proteins have an extended C-terminal sequence following the last cystatin domain, a C-terminal sequence which could arguably be a precursor to a biologically active peptide.

It is also highly probable that a human analogue of Spp24 exists in bone. Specifically, α_2HS glycoprotein, the human analogue of fetuin, circulates in blood as a two chain molecule and that, like Spp24, the cleavages which generate the two chain form occur within the extended C-terminal sequence that follows the last cystatin domain. The connecting peptide which is removed by those cleavages contains a sequence which could be phosphorylated by the SXE/S(P)-specific secretory pathway protein kinase, the sequence SPSGE (residues 310-314) in the protein.

Further, the phosphorylation of serine residues in Spp24 follows the recognition motif for serine phosphorylation that has been found in most secreted phosphoproteins, including MGP and osteopontin. All but one of the serine residues phosphorylated in Spp24 have the negatively charged side chain of

glutamate or phosphoserine in the n + 2 position in the concensus recognition sequence Ser-X-Glu/Ser(P), a substrate recognition pattern first observed in milk caseins and now identified in a wide variety of secreted phosphoproteins. The only phosphorylated serine that does not conform to this recognition motif is serine 130, which has a glycine residue in the n + 2 position in the sequence SSSSGSSSS.

In addition, phosphoproteins secreted into the extracellular environment of cells are invariably partially phosphorylated at each target serine residue, while those phosphoproteins secreted into milk and saliva are fully phosphorylated. This pattern of partial serine phosphorylation is also seen in Spp24. Such partial serine phosphorylation may reflect a role for phosphoserine residues in the regulation of phosphoprotein activity by modulation of the SXE/S(P) specific protein kinase or of a phosphoprotein phosphatase. Many secreted phosphoproteins have phosphoserine residues that are clustered in highly anionic sequences, including MGP and osteopontin. Spp24 also follows this pattern, with a potential maximum net negative charge of 18 in an 11 residue span, assuming complete phosphorylation of all serine residues.

EXAMPLE VI
IDENTIFICATION OF h-Spp24 IN A HUMAN cDNA LIBRARY

A cystatin region of the polynucleotide encoding bovine Spp24 protein (see, SEQ.ID.No.7, FIGURE 5 [from nucleotides 151 to 527]; and, Spp24-mid [ATCC Designation No. 97012]) was labelled with ^{32}P and utilized as a probe to screen a human liver λgt11 library (CLONTECH, Palo Alto, CA). Bacteriophages were plated and transferred to nitrocellulose filters as described by the Clontech protocol. cDNA inserts of these clones were enzymatically amplified by performing PCR with a specific antisense primer and either a λg11 forward or a λgt11 reverse primer (PROMEGA, Madison, WI). An 800 bp clone was

obtained and identified as the full-length cDNA human Spp24 product by PCR hybridization using the 3' and 5' termini of bovine Spp24 (SEQ. ID. No. 7). The human Spp24 clone is available from the American Type Culture Collection under ATCC Designation No. 69745 (deposited January 20, 1995).

5

SUMMARY OF SEQUENCES

SEQUENCE ID. NO. 1 is an N-terminal sense degenerate primer for Spp24.

SEQUENCE ID. NO. 2 is an internal antisense degenerate primer for Spp24.

SEQUENCE ID. NO. 3 is a specific internal sense primer for Spp24.

SEQUENCE ID. NO. 4 is a 23-base oligonucleotide adapter primer for Spp24 linked to a 17-oligo(dt) tag.

SEQUENCE ID. NO. 5 is an adapter primer for Spp24.

SEQUENCE ID. NO. 6 is a specific antisense primer for Spp24.

SEQUENCE ID. NO. 7 is the nucleotide sequence for the Spp24 cDNA.

SEQUENCE ID. NO. 8 is the deduced amino acid sequence for the Spp24 protein.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

(ii) TITLE OF THE INVENTION: NOVEL PHOSPHOPROTEIN SECRETED IN THE EXTRACELULAR MATRICES OF MAMMALIAN ORGANS AND METHODS FOR USE THEREOF

(iii) NUMBER OF SEQUENCES: 8

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Robbins, Berliner & Carson
(B) STREET: 201 N. Figueroa Street, 5th Floor
(C) CITY: Los Angeles
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 90012-2628

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Berliner, Robert
(B) REGISTRATION NUMBER: 20,121
(C) REFERENCE/DOCKET NUMBER: 5555-351

(viii) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 213-977-1001
(B) TELEFAX: 213-977-1003
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base
(B) LOCATION: 6...6
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY: Modified Base
(B) LOCATION: 9...9
(D) OTHER INFORMATION: Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTTCCNGTNT ATGATTATGA 20

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base
(B) LOCATION: 3...3
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY: Modified Base
(B) LOCATION: 9...9
(D) OTHER INFORMATION: Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AANATATCNC CAAAAAACAT 20

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGCTGCCACT GGTCCTCCAG CTCT 24

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ACGCGTCGAC CTCGAGATCG ATG 23

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ACGCGTCGAC CTCGAGATCG ATG 23

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CAGATAGGGG CTCATGACTG GGA 23

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 829 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 92...691
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
ACAGTCTGAT CTGCCAAGTG CATTATACCA ATATCTCATT AATTCTCCCC AAACCTCTGA     60
ACGGAAATTG TTCTTCCCAT TAATGGAGAA G ATG GCG ATG AAG ATG TTG GTG       112
                                   Met Ala Met Lys Met Leu Val
                                    1               5

ATA TTT GTC CTT GGA ATG AAC CAC TGG ACT TGT ACA GGT TTC CCG GTG      160
Ile Phe Val Leu Gly Met Asn His Trp Thr Cys Thr Gly Phe Pro Val
        10                  15                  20
```

```
TAT GAC TAT GAC CCG GCT TCC CTG AAG GAG GCT CTC AGC GCC TCT GTG      208
Tyr Asp Tyr Asp Pro Ala Ser Leu Lys Glu Ala Leu Ser Ala Ser Val
    25                  30                  35

GCA AAA CTG AAT TCC CAG TCA CTG AGC CCC TAT CTG TTT CGG GCG TTT      256
Ala Lys Leu Asn Ser Gln Ser Leu Ser Pro Tyr Leu Phe Arg Ala Phe
40                  45                  50                  55

AGA AGC TCA GTT AAA AGA GTC AAC GCC CTG GAC GAG GAC AGC TTG ACC      304
Arg Ser Ser Val Lys Arg Val Asn Ala Leu Asp Glu Asp Ser Leu Thr
                60                  65                  70

ATG GAC TTA GAG TTC AGG ATT CAA GAG ACG ACG TGC AGG AGG GAA TCT      352
Met Asp Leu Glu Phe Arg Ile Gln Glu Thr Thr Cys Arg Arg Glu Ser
            75                  80                  85

GAG GCA GAC CCC GCC ACC TGT CAC TTC CAG AGG GGC TAC CAC GTG CCC      400
Glu Ala Asp Pro Ala Thr Cys His Phe Gln Arg Gly Tyr His Val Pro
        90                  95                  100

GTG GCC GTT TGC AGA AGC ACC GTG CGG ATG TCT GCT GAA CAG GTG CAG      448
Val Ala Val Cys Arg Ser Thr Val Arg Met Ser Ala Glu Gln Val Gln
    105                 110                 115

AAC GTG TGG GTT CGC TGC CAC TGG TCC TCC AGC TCT GGG TCC AGC AGC      496
Asn Val Trp Val Arg Cys His Trp Ser Ser Ser Ser Gly Ser Ser Ser
120                 125                 130                 135

AGT GAA GAG ATG TTT TTT GGG GAT ATC TTG GGA TCC TCT ACA TCA AGA      544
Ser Glu Glu Met Phe Phe Gly Asp Ile Leu Gly Ser Ser Thr Ser Arg
                140                 145                 150

AAC AGT TAC CTG CTT GGC CTC ACT CCT GAC AGA TCC AGA GGT GAA CCA      592
Asn Ser Tyr Leu Leu Gly Leu Thr Pro Asp Arg Ser Arg Gly Glu Pro
            155                 160                 165

CTT TAT GAA CCA TCA CGT GAG ATG AGA AGA AAC TTT CCT CTT GGA AAT      640
Leu Tyr Glu Pro Ser Arg Glu Met Arg Arg Asn Phe Pro Leu Gly Asn
        170                 175                 180

AGA AGG TAC TCG AAC CCG TGG CCC AGA GCA AGA GTA AAC CCT GGC TTT      688
Arg Arg Tyr Ser Asn Pro Trp Pro Arg Ala Arg Val Asn Pro Gly Phe
    185                 190                 195

GAG TGACACCCTT GAGCAAAATG CACTGGAAGG AATAGAAGTT CCAATGAAGA AAGATAC   748
Glu
200

CTTATGAATT GTGTAATTTT CTTTTGATCA ATTGCAGTCC CTAATAAATG GCTTACTTTT    808
CCTCTTCAAA AAAAAAAAAA A                                              829
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 200 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Ala Met Lys Met Leu Val Ile Phe Val Leu Gly Met Asn His Trp
 1               5                  10                  15
Thr Cys Thr Gly Phe Pro Val Tyr Asp Tyr Asp Pro Ala Ser Leu Lys
            20                  25                  30
Glu Ala Leu Ser Ala Ser Val Ala Lys Leu Asn Ser Gln Ser Leu Ser
        35                  40                  45
Pro Tyr Leu Phe Arg Ala Phe Arg Ser Ser Val Lys Arg Val Asn Ala
    50                  55                  60
Leu Asp Glu Asp Ser Leu Thr Met Asp Leu Glu Phe Arg Ile Gln Glu
65                  70                  75                  80
Thr Thr Cys Arg Arg Glu Ser Glu Ala Asp Pro Ala Thr Cys His Phe
                85                  90                  95
Gln Arg Gly Tyr His Val Pro Val Ala Val Cys Arg Ser Thr Val Arg
            100                 105                 110
Met Ser Ala Glu Gln Val Gln Asn Val Trp Val Arg Cys His Trp Ser
        115                 120                 125
Ser Ser Ser Gly Ser Ser Ser Ser Glu Glu Met Phe Phe Gly Asp Ile
    130                 135                 140
Leu Gly Ser Ser Thr Ser Arg Asn Ser Tyr Leu Leu Gly Leu Thr Pro
145                 150                 155                 160
Asp Arg Ser Arg Gly Glu Pro Leu Tyr Glu Pro Ser Arg Glu Met Arg
                165                 170                 175
Arg Asn Phe Pro Leu Gly Asn Arg Arg Tyr Ser Asn Pro Trp Pro Arg
            180                 185                 190
Ala Arg Val Asn Pro Gly Phe Glu
        195                 200
```

CLAIMS

1. A polynucleotide which encodes a Spp24 polypeptide and fragments thereof which are hybridizable under stringent conditions.

2. The polynucleotide according to Claim 1 having a nucleotide sequence as shown in SEQUENCE ID No.7.

3. A recombinant expression vector containing the polynucleotide of Claim 1.

4. The recombinant expression vector of Claim 3 wherein the vector is the λgt11 phage vector deposited with the American Type Culture Collection as human Spp24 on January 5, 1995.

5. The polynucleotide according to Claim 1 wherein the polynucleotide encodes a human Spp24 polypeptide.

6. Isolated, substantially pure Spp24 polypeptide.

7. The polypeptide of Claim 6 having an amino acid sequence as shown in SEQUENCE ID No. 8.

8. Antibodies which will specifically bind the polypeptide of Claim 6.

9. A pharmaceutical composition containing the polypeptide of Claim 6 and a pharmaceutically acceptable carrier.

10. A method for identifying Spp24 binding molecules comprising:

(a) contacting a candidate molecule suspected of being a Spp24 binding molecule with a Spp24 polypeptide; and,

(b) detecting binding of the candidate molecule by the Spp24 polypeptide or fragment thereof.

11. The method according to Claim 10 wherein the candidate molecule is contained in a biological sample.

12. The method according to Claim 11 wherein the biological sample is of mammalian bone or synovial fluid.

13. A method for detecting connective tissue metabolism in a mammal comprising:

(a) contacting a biological sample of the mammal suspected of containing a Spp24 binding molecule with a Spp24 polypeptide or a fragment thereof; and,

(b) detecting binding of Spp24 binding molecules by the Spp24 polypeptide in the sample.

14. The method according to Claim 13 wherein the connective tissue of interest is mammalian bone.

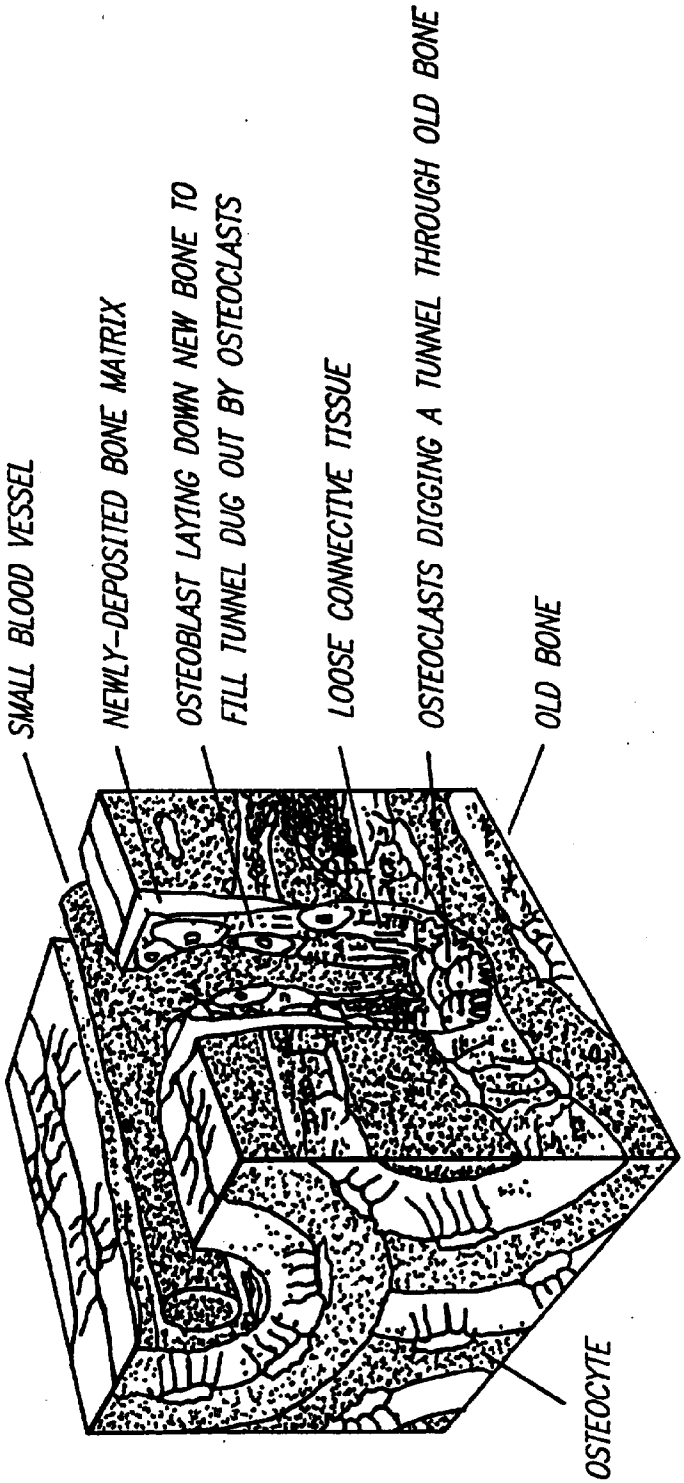
15. A method for detecting connective tissue metabolism in a mammal comprising:

(a) contacting a biological sample suspected of containing Spp24 with an antibody to Spp24 polypeptide; and,

(b) detecting binding of Spp24 polypeptide by the antibody in the sample.

16. The method according to Claim 15 wherein the connective tissue of interest is mammalian bone.

FIG. 1



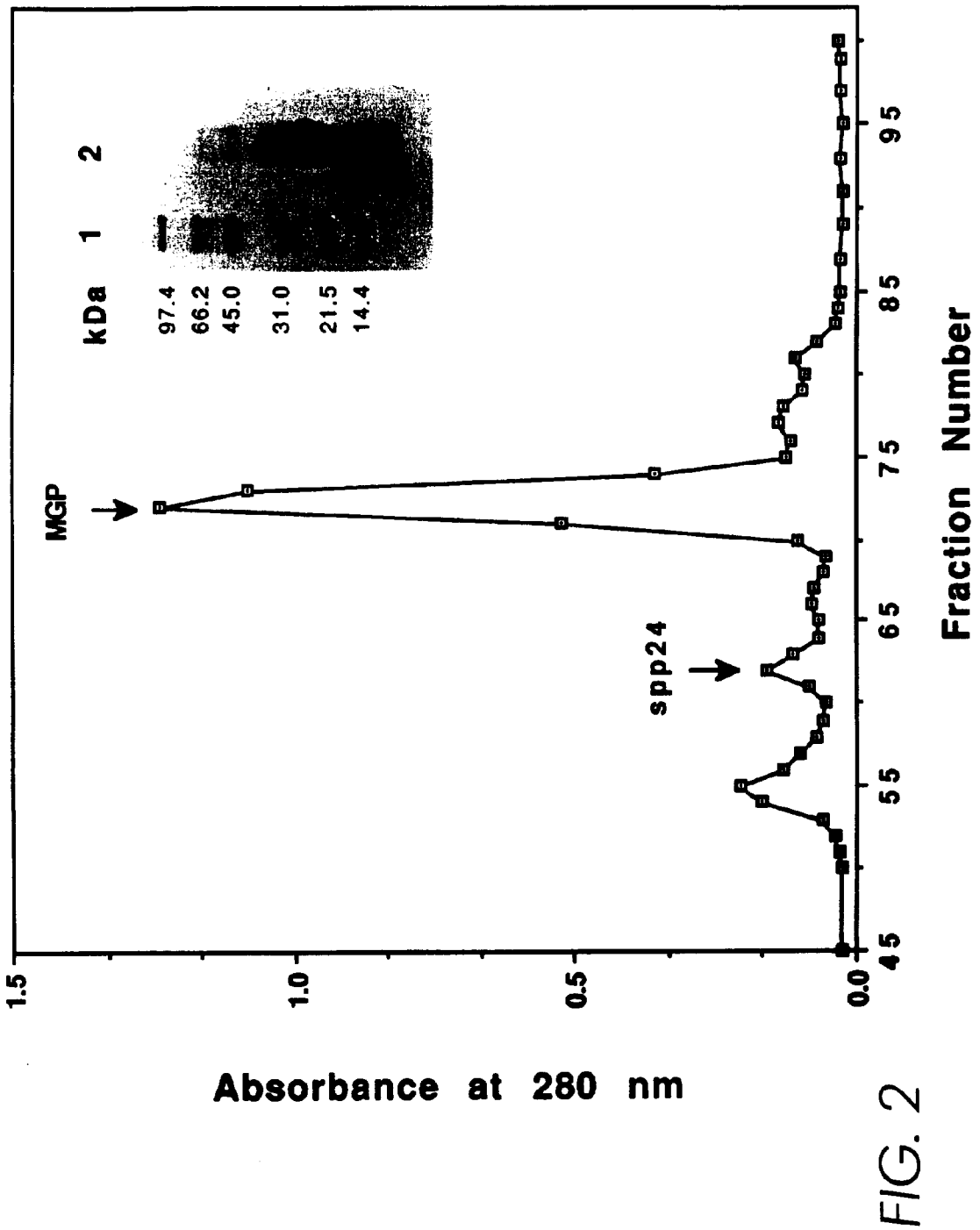

kDa 1 2
97.4
66.2
45.0
31.0
21.5
14.4
MGP
spp24
1.5
1.0
0.5
0.0
45
55
65
75
85
95
Absorbance at 280 nm
Fraction Number

FIG. 2

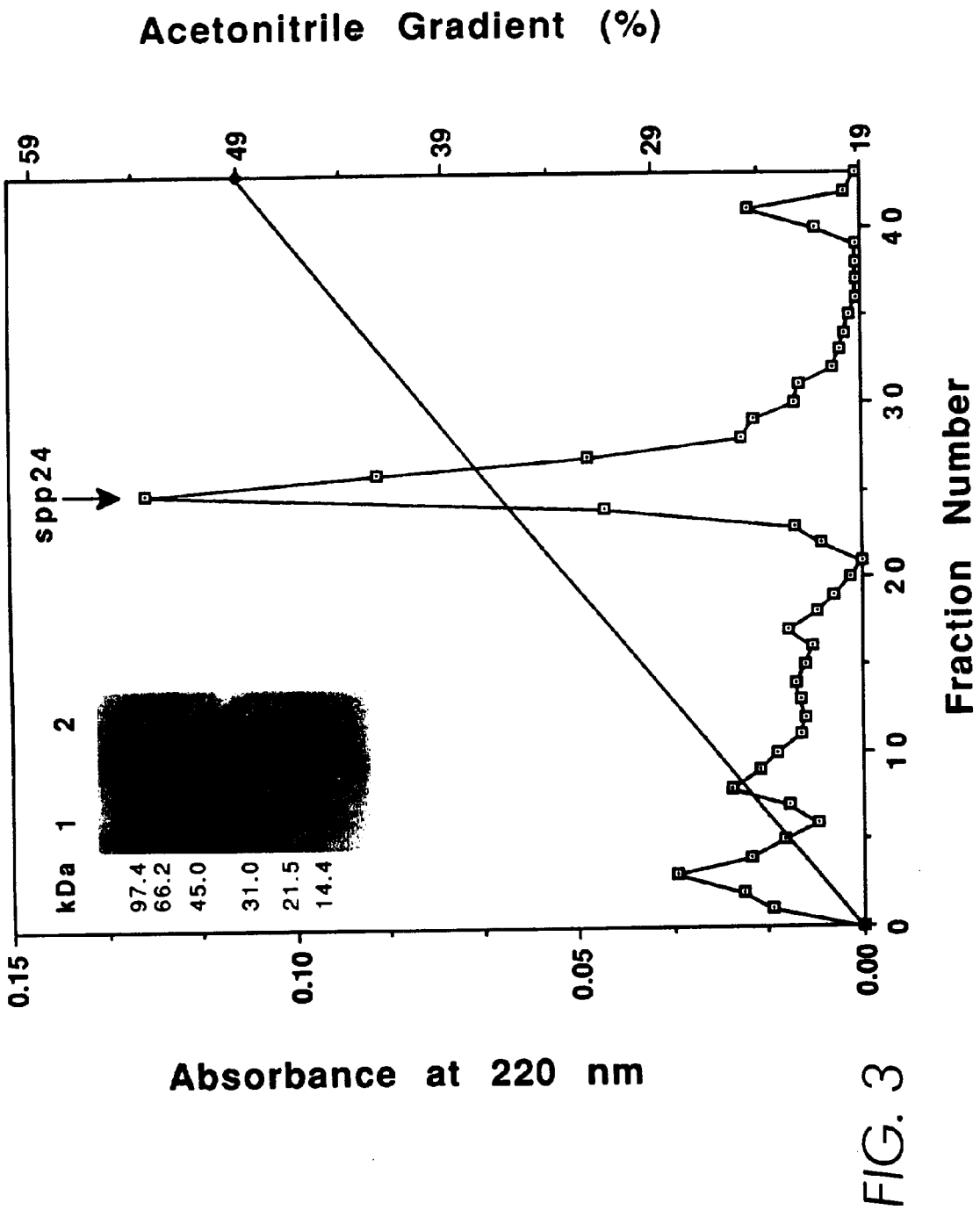

Acetonitrile Gradient (%)
59
49
39
29
19
spp24
kDa 1 2
97.4
66.2
45.0
31.0
21.5
14.4
0.15
0.10
0.05
0.00
0
10
20
30
40
Fraction Number
Absorbance at 220 nm

FIG. 3

FIG. 4

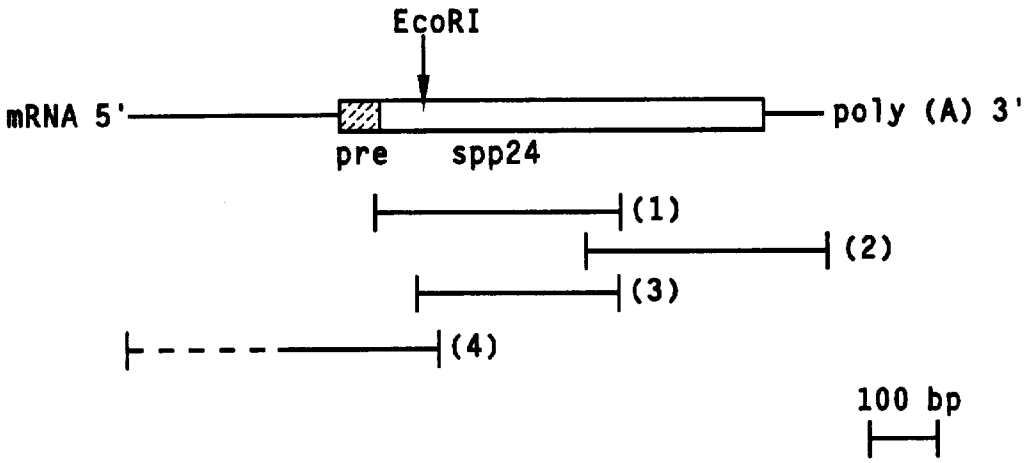

EcoRI
mRNA 5'
poly (A) 3'
pre
spp24
(1)
(2)
(3)
(4)
100 bp

FIG. 6

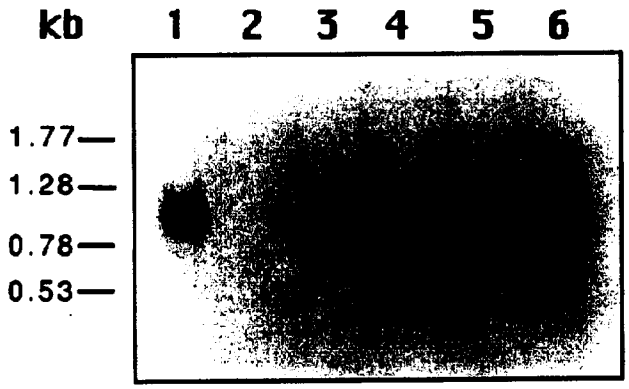
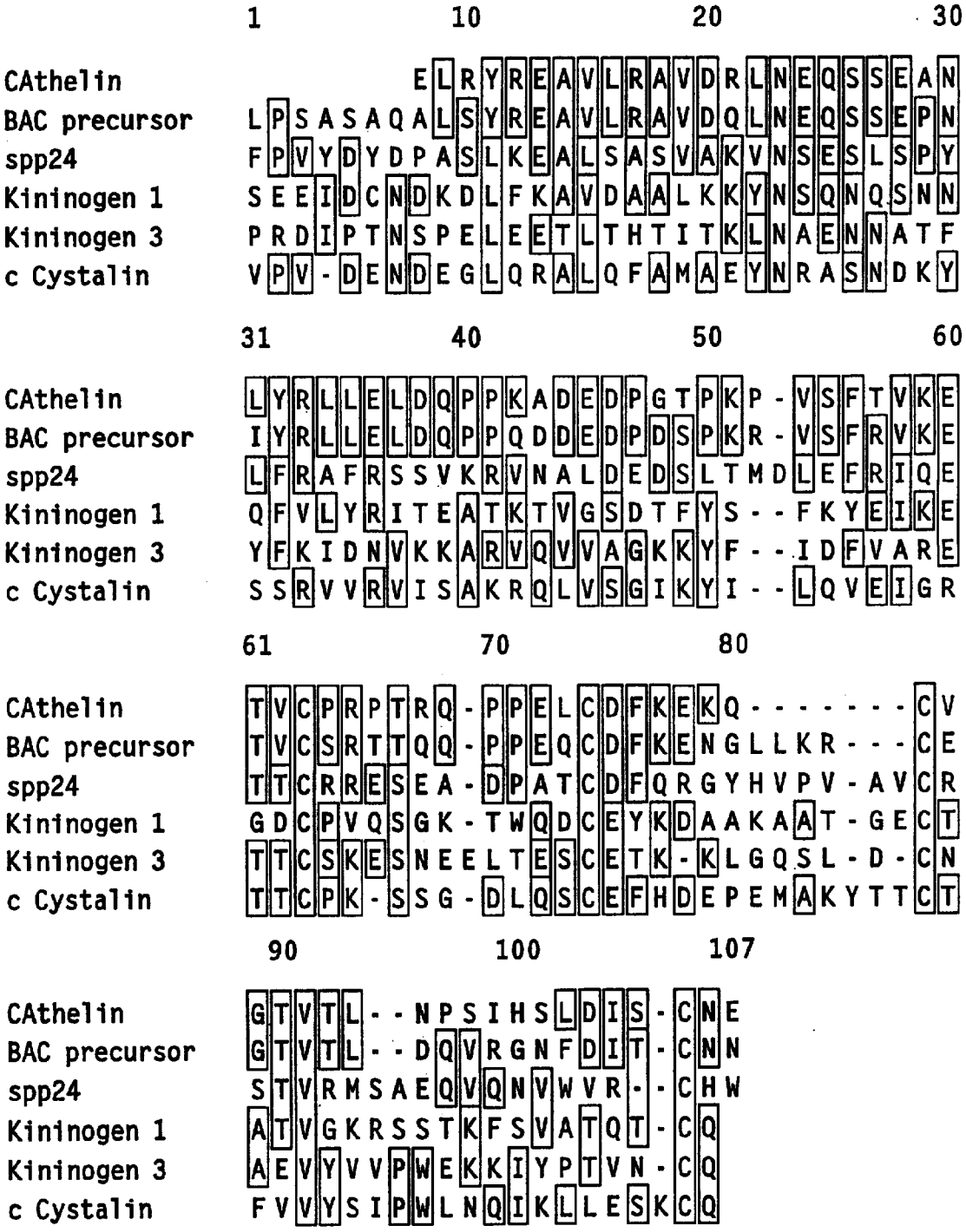

kb
1 2 3 4 5 6
1.77
1.28
0.78
0.53

FIG. 5

```
   1 ACAGTCTGATCTGCCAAGTGCATTATACCAATATCTCATTAATTCTCCCCAAACCTCTGA  60
  61 ACGGAAATTGTTCTTCCCATAATGGAGAAGATGGCGATGAAGATGTTGGTGATATTTGTC 120
                                    MetAlaMetLysMetLeuValIlePheVal  10

 121 CTTGGAATGAACCACTGGACTTGTACAGGTTTCCCGGTGTATGACTATGACCCGGCTTCC 180
  11 LeuGlyMetAsnHisTrpThrCysThrGlyPheProValTyrAspTyrAspProAlaSer  30

 181 CTGAAGGAGGCTCTCAGCGCCTCTGTGGCAAAACTCAATTCCCAGTCACTGAGCCCCTAT 240
  31 LeuLysGluAlaLeuSerAlaSerValAlaLysValAsnSerGlnSerLeuSerProTyr  50

 241 CTGTTTCGGGCGTTTAGAAGCTCAGTTAAAAGAGTCAACGCCCTGGACGAGGACAGCTTG 300
  51 LeuPheArgAlaPheArgSerSerValLysArgValAsnAlaLeuAspGluAspSerLeu  70

 301 ACCATGGACTTAGAGTTCAGGATTCAAGAGACGACGTGCAGGAGGGAATCTGAGGCAGAC 360
  71 ThrMetAspLeuGluPheArgIleGlnGluThrThrCysArgArgGluSerGluAlaAsp  90

 361 CCCGCCACCTGTGACTTCCAGAGGGGCTACCACGTGCCCGTGGCCGTTTGCAGAAGCACC 420
  91 PluAlaThrCysAspPheGlnArgGlyTyrAlsValProValAlaValCysArgSerThr 110

 421 GTGCGGATGTCTGCTGAACAGGTGCAGAACGTGTGGGTTCGCTGCCACTGGTCCTCCAGC 480
 111 ValArgMetSerAlaGluGlnValGlnAsnValTrpValArgCysHisTrpSerSerSer 130

 481 TCTGGGTCCAGCAGCAGTGAAGAGATGTTTTTTGGGGATATCTTGGGATCCTCTACATCA 540
 131 SerGlySerSerSerSerGluGluMetPhePheGlyAspIleLeuGlySerSerThrSer 150

 541 AGAAACAGTTACCTGCTTGGCCTCACTCCTGACAGATCCAGAGGTGAACCACTTTATGAA 600
 151 ArgAsnSerTyrLeuLeuGlyLeuThrProAspArgSerArgGlyGluProLeuTyrGlu 170

 601 CCATCACGTGAGATGAGAAGAAACTTTCCTCTTGGAAATAGAAGGTACTCGAACCCGTGG 660
 171 ProSerArgGluMetArgArgAsnPheProLeuGlyAsnArgArgTyrSerAsnProTrp 190

 661 CCCAGAGCAAGAGTAAACCCTGGCTTTGAGTGACACCCTTGAGCAAAATGCACTGGAAGG 720
 191 ProArgAlaArgValAsnProGlyPheGlu***       ***

 721 AATAGAAGTTCCAATGAAGAAAGATACCTTATGAATTGTGTAATTTTCTTTTGATCAATT 780
 781 GCAGTCCCTAATAAATGGCTTACTTTTCCTCTTCAAAAAAAAAAAAAA              829
```

FIG. 7



SUBSTITUTE SHEET (RULE 26)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/15796

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C12N 7/01, 15/12, 15/70

US CL : 435/69.1, 69.4, 320.1, 235.1; 536/23.5; 935/9, 27, 31, 55

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 69.4, 320.1, 235.1; 530/350; 536/23.5; 935/9, 27, 31, 55

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	EP, A, 0 409 472 A1 (CHIRON CORPORATION) 23 January 1991, Figures 3 and 5, examples 5-6 and 8-9 on pages 14-17.	1, 3, 5 --------- 2, 4
Y	Biochemistry Biophysics Research Communications, Volume 168, Number 3, issued 1990, MATSUURA et al., "Characterization of a 5'-Flanking Region of the Human Liver/Bone/Kidney Alkaline Phasphatase Gene: Two Kinds of mRNA from a Single Gene", pages 993-1000, especially pages 993 and 994, lines 6-10.	2
Y	WATSON et al. Molecular Biology of the Gene. California: Benjamin/Cummings Publishing Company, Inc. 1987, Volume 1, pages 210 to 212.	4

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 21 MARCH 1996	Date of mailing of the international search report 22 APR 1996
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer CHRISTINE SAOUD Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/15796

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 4,788,135 (DAVIS et al.) 29 November 1988, column 10, lines 18-32.	4
A	US, A, 4,804,744 (SEN) 14 February 1989, column 8, lines 10-28.	1-5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/15796

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-5

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/15796

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, STN, BIOSIS, MEDLINE, EMBASE, WPIDS

search terms: bone morphogen?, protein?, bone morphogenetic protein, cysteine protease inhibitor, b :p, cysteine protease, cdna

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

Group I, claims 1-5, drawn to DNA encoding Spp24.
Group .I, claims 6-7 and 9, drawn to Spp24 protein.
Group III, claim 8, drawn to an antibody to Spp24 protein.
Group IV, claims 10-12, drawn to an assay for Spp24.
Group V, claims 13-16, drawn to an assay for detecting tissue metabolism by identification of Spp24.

The inventions listed as Groups I-V do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the inventions lack a special technical feature. The feature common to all the claims, Spp24 protein, was known in the prior art as BMP or bone morphogenetic protein. Special technical features are those features that define over the prior art. In the instant application, since Spp24 was already known in the prior art, this feature does not define over the prior art and unity of invention is lacking.